

SOME REMARKS ON STATISTICAL EDUCATION

BU-537-M

by

March 1978
(Revised)

Walter T. Federer*

Abstract

A definition of the subject Statistics is given first and then the difference between the chalkboard world of the teacher of Statistics as opposed to the real world of the experimenter is stressed. An over-emphasis on significance testing, hypothesis testing, and decision procedures has led to a de-emphasis of statistical design. The teaching of statistical design theory, the teaching of Statistics in a changing world, the importance of model building, and different approaches to teaching Statistics are discussed. Some published materials developed to meet teaching needs and a new type of Statistics course are described. Information about special issues on Statistical Education (teaching and consulting) is presented.

KEY WORDS: Teaching innovations; Statistical design; Survey design; Model building; Hypothesis testing; Special publications on statistical education.

1. Introduction

In order to ensure that we are talking about the same topic, a definition of the subject STATISTICS is presented. One frequently used is that Statistics is concerned with the characterization, development, and application of techniques for

- (i) The statistical design of an investigation whether it be an experiment, a survey, an observational, or a model building study,
- (ii) The summarization of the facts from the investigation, and
- (iii) The inferences that can be drawn from the facts of the investigation about the parameters in the population (or generalizing from the specific sample to the general population).

* Walter T. Federer is Professor of Biological Statistics in the Biometrics Unit, Cornell University, Ithaca, New York 14853.

It is my view that the statistical design part of the definition of Statistics encompasses the following items (see Chapter 4 of [3] for definitions of terms): Description of Variables and Populations, Measurements and Measuring Instruments, Treatment Design, Experiment Design, Sequential Design, Sample Survey Design, Model Building Design, Determination of Sample Size, and Principles and Properties of Statistical Designs. These items must be considered for either planned or unplanned investigations.

With regard to the second part of the definition of Statistics, this is and has been an important research and teaching area for statisticians. Considerable effort is expended in developing and in applying statistical procedures for summarizing facts from a set of data obtained from an investigation whether planned or unplanned.

With respect to the third part of the definition of Statistics, that is making inferences about population parameters based on sample facts, a considerable amount of effort in teaching and research has been devoted to this topic, almost to the exclusion of the first two parts of the definition in some courses. For example, some teachers tell their classes that "Statistics is inference".

In addition to defining Statistics, the teacher of Statistics, whatever his departmental affiliation, should point out when he is talking about the chalkboard world of the classroom and when he is talking about real world situations. In the former case his assumptions hold, whereas they may not in the latter. It would appear that some teachers of statistics courses never get out of their chalkboard world if one is to judge by their performance. They define the sample obtained to be a random sample from a prescribed population with a known distribution when in fact the particular classroom sample used may be neither random nor from the distribution stated. In other words, statisticians may profess to be in the real world of experimentation when they are really only in their chalkboard world.

Assumptions involving real world data must be tested but assumptions involving classroom examples are obtained by definition. These distinctions must be emphasized for users of statistical procedure.

2. The Over-Emphasis on Significance and Hypothesis Testing

The work of Fisher, Neyman, Pearson, Wald, Jeffreys, and others since the early 1930's has had a pronounced effect on the role of significance testing, hypothesis testing, and decision making in statistical research and teaching. One effect has been to cause the emphasis to be so one-sided that some courses and some books are devoted almost entirely to significance and/or hypothesis testing. In addition, some editors and major professors will accept only results that are significant at the five percent level. (We may deplore this but after all, it is the statisticians' fault that they react in this manner.) There are statistics books which consider that estimation is a part of hypothesis testing, and hence it can be largely ignored. But for non-Bayesians, what about estimation in sequential sampling and experimentation? Another casualty of the statistics teacher's obsession with significance testing, hypothesis testing, and decision making has been a consideration of the statistical design as it affects the other two parts of the definition of Statistics. Many statisticians and teachers of Statistics assume, but do not verify, that they have a random sample from some prescribed population.

Another emphasis in teaching that also de-emphasizes the role of statistical design is statistical computing both with numbers and with symbols. Number-bashing and symbol-bashing are easy items to teach and hence get taught. Many individuals' obsession with a high speed computer has led to emphasis on number-bashing, and those interested in mathematical manipulations tend toward an emphasis on symbol-bashing. All of the above are important, but none is all-important.

When the statistics teacher confines himself to the chalkboard world of the classroom, the sampling procedure need not be verified but can be proclaimed by definition and passed off with the phrase "I have a random sample from my population" or with the phrase "the sample comes from the population". This procedure allows statisticians to ignore the first part of the definition of Statistics. In addition, to a large proportion of statisticians, at least as judged by writings and teachings, experiment design means the statistical analysis of completely randomized, randomized complete block, latin square, balanced incomplete block, split-plot, or other experiment designs. To another more mathematically oriented group, statistical design is a branch of combinatorics, geometry, or topology. These groups have little or no feeling for such concepts as population structure, experimental and sampling units, blocking, randomization, nature of competition, confounding, or replication as practiced by the experimenter. The realization that there are several definitions of orthogonality, balance, connectedness, etc. appears to be lost on many. The fact that confounding is almost universal in human and animal experiments would come as a surprise to many in these groups even if they understood confounding concepts.

From a livelihood and user's viewpoint, perhaps the most surprising aspect of today's teaching of Statistics is the lack of sample survey design courses in biology, in business, in industry, in health studies, etc. and of course in questionnaire construction, interviewing techniques, quantification, and measurements. We should note that a large proportion of the world's statisticians make their living designing sample surveys and summarizing data from these surveys. One has only to think of the large number of statisticians employed in the many and diverse federal bureaus of the U. S. such as the Census Bureau, the Bureau of Labor Statistics, etc. In addition, when one considers private survey, state government, and international organizations, the list of statisticians employed in designing and summarizing

sample surveys becomes very large. The course offerings in Statistics at universities and colleges tend to ignore this large group and to concentrate their efforts on reproducing replicas of themselves (possibly "in the limit"). Even the prestigious CUPM Panel on Statistics (F. Graybill, Chairman, R. Bradley, H. Chernoff, P. Clifford, S. Goldberg, J. Neter, and G. Nicholson) which produced the CUPM Bulletin, does not include statistical design in their recommended key courses. Experiment and survey design are considered to be additional topics on a parallel with robustness, factor analysis, nonparametric statistics, etc. which may be selected as additional topics or courses depending upon the interest of the instructor. Of course, this is very much in line with statistical offerings at universities and colleges in that a course on sample survey design is treated like a step-child and a course on interviewing techniques as a very infrequent, unwanted visitor! How many of the statistical centers listed in The American Statistician every year can boast of a professor whose main interest is in the area of survey design and interviewing techniques? Paul G. Homeyer, Research Triangle Institute, recently stated that "it is very difficult to find sample survey specialists among new Ph.D. graduates in Statistics, and we frequently select from available university graduates and then have to train them in the theory and application of survey design".

A similar situation occurs in bioassay and in sequential experimentation. Both are widespread in practice, but applied courses in these two areas are not regular offerings at most universities. Statistical procedures for dosage response and efficiency and efficacy of vaccines are not as advanced as required for present day needs. Standards for drug acceptance and for determining drug side effects are in a sad state. Studies of statistical procedures on the use and management of technology in society do not receive funding or attention in today's world. Also, a Snedecor-Cochran type text using sequential procedures and random sample sizes is not to be found. One should note that much of experimentation is sequential

in nature and that random sample sizes occur frequently in investigations, e.g., post-stratification in surveys, unequal numbers n-way classifications, etc.

3. Teaching Statistics in a Changing World

The world of science is a constantly changing phenomenon with new technology, for example the high speed computer, molecular chemistry, single-cell culturing, cloning, etc., often demanding new techniques and procedures in areas such as Statistics. If statisticians do not move to meet these demands via research and teaching, others will move in to fill the void. Such is the situation with regard to model building in the world of science. Scientists from many diverse areas are filling the void in building deterministic models but are doing relatively little in the area of stochastic model building. Perhaps the lack of statistical interest in model building is because statisticians really believe in a defined linear model world, they believe only in the traditional approach in Statistics, or they do not see the problem. Model building is and should be a part of Statistics. Will statisticians let model building become established in other areas or as another new area outside of Statistics? Or, will they make it a part of Statistics?

Traditionally, courses in Mathematics and Statistics have tended to become stereotyped. In Statistics, there is a tendency for recently trained theoretical statisticians to utilize a probability stereotype approach in teaching an introductory probability and mathematical statistics course or a statistical methods course. The latter type of course often will rely heavily on the probability content with numerical manipulations substituted for the mathematical manipulations. A "nonmathematical" introductory course may essentially be a "watered-down" version of an introductory probability and mathematical statistics course, resulting in a probability stereotype. Also, in such courses the sample will be defined to come from a specified population (usually normal) and hence there is no need for teach-

ing statistical design. The resulting data will be assumed to be informative and collected in a manner to produce a "representative" sample from a population. Methods of ascertaining the representativeness of a sample, the reliability of measurements and measuring devices, the appropriateness of an assumed model, or the validity of applying a statistical procedure receive short shrift in the teaching of Statistics.

Several approaches to the teaching of Statistics are available. For example, there is the probability and inferential approach, there is the methodological or data summarization and inferential approach, and there is a data analysis approach as presented by John Tukey. These approaches constitute those used in the majority of statistical texts, illustrating the lack of emphasis devoted to the first part of the definition of Statistics. A logical pedagogical approach would be to present Statistics as it is defined with the order (i), (ii), and (iii) for the three parts of the definition rather than the order (ii) and (iii) or (iii) and (ii), with (i) being part of a definition as represented in current statistical methods and mathematical statistics texts.

In any event, the teaching of Statistics should not be confined to the probability-stereotype as Mathematics has been to the calculus-stereotype. Rather, several new approaches should be tried with one of them being model building. Another approach might be to team up with a History of Science person to study the development of Statistics as affected by the political, social, and scientific climate of the times. Another approach, to be described later, would be to relate concepts of, reasons for, and methods of collecting and summarizing meaningful and reliable data with much less emphasis on inference than present courses. This type of approach would be closer to what the vast majority of government statisticians are now doing. Additional ideas and comments may be found in papers by Blyth [1], Good [5], and Hogg [7].

4. Development of Statistical Materials to Meet Needs

One approach taken to meet some of the statistical needs of the real world as opposed to the chalkboard world of the professor has been to develop a number of case studies of applications and applicability of statistical techniques. Some materials the author has found useful in this area are found in references [9], [11], [14], [15], and [18]. Although these materials were found useful, it was not possible to use them as textbooks for a course in Statistics. It is felt, however, that every teacher of Statistics should have access to them.

Another approach to the teaching of Statistics was initiated in the Spring of 1966 following discussions in the Spring of 1965 on undergraduate offerings in Statistics as related to the needs of all students at Cornell University. It was decided that a completely different course from the usual statistical methods or mathematical statistics and probability courses was needed for some first and second-year college students. In particular, a course presenting the basic ideas, concepts, and philosophies associated with the measurement and procurement of numbers in investigations and of methods of data summarization was needed. A minimum of emphasis on statistical methodology or on mathematical manipulations was required for this group of students. The course would emphasize the relationship of numbers and data to many aspects of a society. Examples would be current from newspapers, magazines, and textbooks, and data collected by the students would be utilized. The course was to be complementary to and mostly nonrepetitive of other introductory courses on Campus. It was to be an "artsie" type course in Statistics in that the student would learn about the subject of Statistics and statistical ideas but would not attain proficiency in the technical aspects of the subject. Students were to be made aware of how Statistics (the subject) and various statistics affect their everyday lives. The course title adopted was:

STATISTICS AND THE WORLD WE LIVE IN

and a text was written for this course (see [3] and [4]). The topics in the course include:

1. An introduction to Statistics and how it affects society.
2. Measurements, measuring, variability, bias, precision, accuracy, and a linear model composed of assignable causes plus non-assignable causes.
3. A class survey and data collection.
4. The what, where, why, who, and how of data collection.
5. Phony statistics, with numerous examples from the news media, politicians' speeches, advertisements, etc.
6. The principles of scientific experimentation.
7. Sample survey designs, questionnaire construction, randomized and block total response.
8. Experiment designs and properties of designs.
9. Treatment designs and their properties.
10. Graphs, charts, stem and leaf diagrams, and tabular arrays.
11. Elementary probability.
12. Measures of location and dispersion for various sample survey, experiment, and treatment designs.
13. Organized or patterned variation with the binomial, Poisson, and normal distributions as examples and confidence interval estimation.
14. Sample size determinations.
15. Some U.S. and U.N. statistics and statistical publications.
16. A further introduction to statistical methodology but involving only concepts presented previously. Correlation, regression, analysis of variance, and t and χ^2 tests of significance used as summary statistics, are discussed. Hypothesis testing is not considered.

Wallis and Roberts [16] is useful in supplementing the material for items 1, 5, and 10-13. Wilson [17] is used to supplement topics 2, 3, 4, and 6 above. Campbell [2], Huff [8], and the first paper, by D. Seligman, in Sielaff [11] are being used as sources for examples of phony statistics. However, one has to be careful to dispel the negative aspects of Statistics created by presenting phony statistics. Slonim [12] is used to supplement text material for item 7 above.

Two additional features of the course are the taking of a class survey, as described in Section II.3, and the independent sample survey project described in problem V.2, page 84, of [3]. It has been found that groups of students, say 2 to 4, obtain much more reliable and meaningful results than do students working alone. These group projects are a highlight of the course. From the approximately 25 projects carried out by over 90 students in Spring 1974, two papers have been published in professional journals, several more came under consideration for publication in campus publications, one resulted in a new statistic in the area of randomized response, and one is a possibility for a master's thesis in physical education, should the student continue in that area. Topics such as attitudes toward and information about abortion, birth control, campus bus service, streaking, race horse breeding, care of flowers by registered New York florists, dogs on campus, temperatures in buildings, gardening, baseball batting averages, and student housing were among those studied. It is amazing how much time students will devote to a project which they select, which interests them, and when they are working with someone they enjoy. In Fall 1973 one of the studies resulted in one campus library opening for additional hours in order to accommodate student desires. Students can observe that good reliable information, handled in an appropriate manner, can effect changes. Also, they learn to develop a questionnaire, to plan and conduct a survey, and to write a report on the survey results. Some students elect to conduct an experiment rather than a survey.

Also, some emphasis is given to interviewing techniques such as the randomized response technique and the block total response procedures (see Smith, Federer, and Raghavarao [13]). These are used on students in one of the weekly discussion periods. Several groups of students used the randomized response procedure in Spring 1974 in their group projects. Some time in the course is devoted to writing and coding questionnaires, as well as to development of sampling frames and selection of sampling units.

Three types of introductory courses have been taught at Cornell University since 1966, i.e., mathematics and probability, statistical methodology, and statistical concepts and ideas in relation to society. Recently, an introductory statistics course which relies heavily on data analysis has been presented. This gives four types of introductory courses on campus. At least two courses on deterministic model building are being taught, one in Agricultural Engineering and one in Electrical Engineering. A course on stochastic modeling, with mathematics and probability prerequisites, was first given in Spring 1976. This diversity of offerings meets more needs of the students and more areas of interest. Also, in experimenting with courses and topics, it appears that the modular approach may have merit (see [6], e.g.). The second term of a Snedecor-Cochran type statistical methods course at Cornell University has been supplemented and subdivided into six modules (short courses) as follows (see [6]):

Part of term	Topic	Topic
1 st one-third	Design and Analysis I	Sampling Biological Populations
2 nd one-third	Design and Analysis II	Regression I
last one-third	Nonparametric and Distribution-Free Statistical Methods	Regression II

Design and Analysis I and part of II and Regression I and part of II constituted the previous second semester course of statistical methods. The remaining material listed above constitutes supplementary material. Two parallel modules were presented in each one-third of the term. Students were allowed to register for one, two, ~~two~~, or six modules. Some advantages of the system are that students are allowed to select modules of interest to them, additional topics in statistics can be presented without adding full-term courses, the modules are ready-made short courses, students can be exposed to several instructors, and more flexibility in student and staff interests and abilities is possible. Some disadvantages encountered in the system are that some students tend to consider one module as a course, additional staff time is required for modules, and an instructor sometimes includes too much material in a module. It now appears that a two-semester sequence of statistical methods supplemented by modules on sampling biological populations, bioassay, distribution-free statistics, life-testing, regression, linear programming, plot technique (for laboratory, greenhouse, and field experiments), statistical plant breeding, etc., would be a more desirable approach than the present one. Students in a module would be more statistically mature in having had two semesters of statistical methods rather than only one. This procedure should eliminate some of the disadvantages listed.

5. Availability of Innovations in Teaching Statistics and in Statistical Consulting

What material, when to include this material in a sequence of courses, and how to include it are challenges presented to teachers of Statistics. These are usually difficult decisions requiring considerable time which is not available to a busy teacher. Then, how will changes occur in the teaching of Statistics? Perhaps one effective method for this would be to establish a journal on statistical

education both in teaching and consulting, e.g., A Proceedings of the Section on Education of the American Statistical Association. The journal should not be too selective, and one would have to realize that useless material might appear. A few good ideas would be worth more than the cost of printing many additional pages of relatively less useful material. Young instructors and consulting statisticians should be encouraged to contribute, possibly before they become molded by traditional approaches. Also, methods of teaching material would be interesting reading for teachers of Statistics. For example, a paper on how Statistics by Example was used in a course or one on how Statistics: A Guide to the Unknown was used as a text in a course, could be extremely helpful to someone else who would like to try the same thing.

Presently there appear to be few places where such articles could be published. The American Statistician has fulfilled a part of this role and is changing its emphasis to meet more of this need. Special publications fulfill another part. For example, a Special Issue of Communications in Statistics (CIS) on Statistical Education, number 10 of volume A5, has been prepared, and a second issue is under consideration. In another endeavor, Råde [10] has edited a book on statistical education. In the book, attention is drawn to a number of related references, and a useful bibliography of source material is included. It is not certain where endeavors of the above nature are heading. The American Statistician may have insufficient space or interest to include many papers like those in Communications in Statistics, A5, number 10, or those in L. Råde's book. (These papers are not of the same nature as those in the Teachers Corner of The American Statistician.) An outlet for such papers is needed, but its nature is still not clear. Whether or not special issues as described above will develop into a full-fledged Statistics teaching journal will depend upon interest and papers. Experimentation with special issues and with present journals should precede any

new journal establishment. The final outcome, supervision, and editorial nature for such a journal need not be determined at the outset, but it is necessary to ascertain the amount of interest in this area as indicated by the volume of papers that would be submitted.

Acknowledgment

The comments of referees, of an Associate Editor, and of the Editor, were very helpful in the preparation of a final version of this article.

References

- [1] Blyth, C. R., "Subjective vs. objective methods in statistics", The American Statistician, 26, No. 3 (1972), 20-27.
- [2] Campbell, S. K., Flaws and Fallacies in Statistical Thinking, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.
- [3] Federer, W. T., Statistics and Society, Marcel Dekker, Inc., New York, 1973.
- [4] Federer, W. T., Statistics and Society Supplement, Marcel Dekker, Inc., New York, 1973.
- [5] Good, I. B., "Statistics and today's problems", The American Statistician, 26, No. 3 (1972), 11-19.
- [6] Guthrie, D., O'Fallon, J. R., and Service, J., Modular Instruction in Statistics, American Statistical Association, Washington, D. C., 1976.
- [7] Hogg, R. V., "On statistical education", The American Statistician, 26, No. 3 (1972), 8-11.
- [8] Huff, D., How to Lie With Statistics, N. W. Norton and Co., Inc., New York, 1954.
- [9] Mosteller, F., Kruskal, W. H., Link, R. F., Pieters, R. S., and Rising, G. R. (Joint Committee on the Curriculum in Statistics and Probability, Amer. Stat. Assoc. and Nat. Council Teachers Math.) Statistics by Example, 4 volumes entitled: Detecting Patterns, Exploring Data, Finding Models, and Weighing Chances, Addison-Wesley Publishing Co., Reading, Mass.; Menlo Park, Calif.; London, England; Don Mills, Ontario, 1973.
- [10] Råde, L. (editor), Statistics at the School Level, Almqvist and Wiksell International, Stockholm, Sweden, p. 242, 1975.

- [11] Sielaff, T. J., Statistics in Action. Readings in Business and Economic Statistics, The Lansford Press, San Jose, California, 1963.
- [12] Slonim, M. J., Sampling (original title was Sampling in a Nutshell), Simon and Schuster, New York, 1960.
- [13] Smith, L. L., Federer, W. T., and Raghavarao, D., "A comparison of three techniques for eliciting answers to sensitive questions", Proceedings of the Social Statistics Section, American Statistical Association, 1974.
- [14] Tanur, J. M. (editor), Mosteller, F., Kruskal, W. H., Link, R. F., Pieters, R. S., and Rising, G. R. (Joint Committee on Curriculum in Statistics and Probability, Amer. Stat. Assoc. and Nat. Council Teachers Math.) Statistics: A Guide to the Unknown, Holden-Day, Inc., San Francisco, Cambridge, London, Amsterdam, 1972.
- [15] Wallis, W. A., Mosteller, F., et al. of the President's Commission on Federal Statistics, Federal Statistics, Volumes I and II, Superintendent of Documents, U. S. Government Printing Office, Washington, D. C., 1971.
- [16] Wallis, W. A. and Roberts, H. V., Statistics: A New Approach, The Free Press, Glencoe, Illinois, 1956.
- [17] Wilson, E. B., An Introduction to Scientific Research, McGraw-Hill Book Co., Inc., New York, London, and Sydney, 1952.
- [18] Zelinka, M. and Weisberg, S., Teacher's Commentary and Solutions Manual for Statistics by Example, 4 volumes entitled: Detecting Patterns, Exploring Data, Finding Models, and Weighing Chances, Addison-Wesley Publishing Co., Reading, Mass.; Menlo Park, Calif.; London, England; Don Mills, Ontario, 1973.

Note: Revised 8/77 and title changed from "Statistics and Society or Statistics and the World We Live In".

8/77 version revised again 3/78.